

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2001, 15:58:17 : Search time 12.89 Seconds

(without alignments)
2030.348 Million cell updates/sec

Title: US-09-587-111-5

Perfect score: 4004

Sequence: 1 MTSPSSSPVRLTLDGGE.....EDEDGASENVVPVQLQSN 764

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	215	5.4	973	1	TRPS_HUMAN
2	198.5	5.0	1275	1	TRP_DROME
3	197.5	4.9	1124	1	TRPL_DROME
4	159	4.0	809	1	TRPL_MOUSE
5	151	3.8	3924	1	ANK2_HUMAN
6	150.5	3.8	793	1	TRP1_HUMAN
7	140	3.5	1862	1	ANK1_MOUSE
8	139	3.5	848	1	TRP3_HUMAN
9	138	3.4	1880	1	ANK1_HUMAN
10	134.5	3.4	1401	1	LATA_LATMA
11	133.5	3.3	642	1	LATA_SCHPO
12	130	3.2	930	1	TRP6_MOUSE
13	129	3.2	1964	1	NTC4_MOUSE
14	128.5	3.2	1093	1	SW14_YEAST
15	124.5	3.1	2703	1	NOTC_DROME
16	123.5	3.1	832	1	ANK3_HUMAN
17	123	3.1	500	1	CACT_DROME
18	121.5	3.0	1503	1	TRP7_HUMAN
19	121	3.0	931	1	TRP6_HUMAN
20	120.5	3.0	741	1	RNSA_HUMAN
21	120.5	3.0	823	1	YNS2_CAEEL
22	119.5	3.0	357	1	SH5A_MOUSE
23	119	3.0	657	1	RES2_SCHPO
24	119	3.0	836	1	TRP3_MOUSE
25	117.5	2.9	613	1	YSV1_CAEEL
26	116	2.9	357	1	SH5A_MOUSE
27	115	2.9	1429	1	L112_CAEEL
28	113	2.8	971	1	KBF1_MOUSE
29	112.5	2.8	806	1	PA26_HUMAN
30	112	2.8	1302	1	MORS_DROME
31	112	2.8	2318	1	NTC3_MOUSE
32	111.5	2.8	764	1	AKR1_YEAST
33	111.5	2.8	1083	1	Y1L2_YEAST

34	111.5	2.8	2444	1	NTC1_HUMAN	P46531 homo sapien
35	111	2.8	637	1	RES1_SCHPO	P33520 schizosacch
36	110.5	2.8	538	1	THRP_HAETIN	P44980 haemophilus
37	110.5	2.8	1431	1	DAK1_HUMAN	P53355 homo sapien
38	109.5	2.7	906	1	KBF2_CHICK	P98150 gallus gall
39	109.5	2.7	2353	1	CCAH_HUMAN	O95180 homo sapien
40	109.5	2.7	2531	1	NTC1_RAT	O07008 ratius norv
41	109	2.7	1873	1	CCAS_RABIT	P07293 oryctolagus
42	108.5	2.7	448	1	NUAM_ANCOU	P33511 anopheles q
43	108.5	2.7	2437	1	NOTC_BRARE	P46530 brachydanio
44	107.5	2.7	614	1	GAAL_YEAST	P39012 saccharomyc
45	107.5	2.7	679	1	RNSA_MOUSE	O05921 mus musculu

ALIGNMENTS

RESULT ID	1	TRPS_HUMAN	STANDARD	PRT	973 AA.
AC	090162:				
DT	01-OCT-2000 (Rel. 40, Created)				
DT	01-OCT-2000 (Rel. 40, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	TRANSIENT RECEPTOR POTENTIAL CHANNEL 5 (HTRP-5).				
GN	TRP5 OR TRP5.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Brain;				
RA	MEDLINE=9425273; PubMed=10493832;				
RA	Soessey-Alaoui K., Lyon J.A., Jones L., Abidi F.E., Hartung A.J.,				
RA	Hane B., Schwartz C.E., Stevenson R.E., Srivastava A.K.;				
RT	"Molecular cloning and characterization of TRP5 (HTRP5), the human				
RT	homologue of a mouse brain receptor-activated capacitative Ca(2+)				
RT	entry channel.";				
RL	Genomics 60:330-340(1999).				
CC	- FUNCTION: SUGGESTED TO MEDIATE CAPACITATIVE CALCIUM ENTRY (CCE).				
CC	- SEEMS TO FORM A CALCIUM PERMEANT CHANNEL.				
CC	- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).				
CC	- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY.				
CC	- SIMILARITY: CONTAINS 2 ANK REPEATS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
CC	EMBL: AF054568; AAF00002.1; -				
DR	Interpro: IPR002110; -				
DR	Interpro: IPR002111; -				
DR	Interpro: IPR002153; -				
DR	Plam: PF00023; ank; 2.				
DR	Ion channel: Transmembrane; Ion transport; Calcium channel;				
KW	ANK repeat; Repeat; Glycoprotein.				
KW	ANK repeat; Repeat; Glycoprotein.				
FT	TRANSMEM 331				
FT	TRANSMEM 364				
FT	TRANSMEM 399				
FT	TRANSMEM 438				
FT	TRANSMEM 471				
FT	TRANSMEM 513				
FT	TRANSMEM 568				
FT	TRANSMEM 604				
FT	REPEAT 69				
FT	REPEAT 141				
FT	REPEAT 170				
FT	CARBOHYD 91				
FT	N-LINKED (GLCNAC. .) (POTENTIAL).				

FT CONFLICT 285 288 GORO -> ASSE (IN REF. 2).
 FT CONFLICT 326 329 RROO -> POE (IN REF. 2).
 FT CONFLICT 365 374 KPEKFTHS -> NPLSSSKRP (IN REF. 2).
 FT CONFLICT 785 785 S -> N (IN REF. 2).
 SQ SEQUENCE 1275 AA; 142589 MW; 91CFCD9896981 CRC64;

Query Match 5.0%; Score 198.5; DB 1; Length 1275;
 Best Local Similarity 19.9%; Pred. No. 2.2e-06;
 Matches 130; Conservative 100; Mismatches 251; Indels 171; Gaps 27;

QY 153 MACSTDYRGHSALHAIERKSLQCVKLVENGANVHARACGRFGKGCCTFYGELP 212
 DB 62 NINCTDPMNR-SALISAIEENFDLWVLEHINIEV-----GDALLHAISEEX 108
 QY 213 LSLAACKQMDVSYLLENPHQASLOADDSQNTV---LHALVMSIDNSAEINAVLTSM 269
 DB 109 VEAVEELLQWEETNH---KEGQPSWEAVDRSKSTFTVDITPLLAHRNNVLEILK1--- 162
 QY 270 YDGLLOAGARL-----CPTVOLED-----TRNLODLPKLAKCKG 306
 DB 163 ---LDRGATLPPHPHYKCGCCDCVTSQTDSLHQSOSRINATRALSASSLALSSRDY 219
 QY 307 IEIFRHILQREFSGLSHLSKRTKFTWCYGPVRSYLDAS--VDSCEENVLEIIFHCKS 364
 DB 220 LTFVFO--LSWELKRLQAMESEFRAE-YTEMROWVODFGTSLDHARTSMELVNLNHE 276
 QY 365 P-----HRHRMVLEP-LNKLQAKWLLLPKF----- 391
 DB 277 PSNDINCLCQROTLEKLAIRYKQTFVANHVPVQDLAATWDGLPGRKQASQOQLMD 336
 QY 392 FLNPLNLTYMFTAVAVYHQPFLKKAAPHLKAEVNSMLTGLHLLILGIGIYLLVQOL 451
 DB 337 VVVLGGSFPIYSKLTLAPDSEGAKFMKRPVKFITHSCSYM--FLMLLGAASLKVQI 394
 QY 452 ---WTF-----WRRH-----VF-----I 461
 DB 395 TPELLAFPMMLTLMEDMRKHERGSLPGLIELATITTYIMALIFEELKSLVSDGLFEYIMDL 454
 QY 462 W--ISFIDSYFELLFLF--QALLTVVSQVLCFLAIEMYL-----PLTVS-----A 502
 DB 455 WNIIVDYISNMFTYTWILCRATANVIYHRDIFWFGIDYPRPREHMPDPDLSEGAFAAG 514
 QY 503 LVLGWNLILYTRGFOHTGIYSVMIQKVIILRDLRFLLYLVLEFGFAVALVSLSEAMR 562
 DB 515 MWFSYLIKIVHIEFINHIGRPLOVSLGRMII-DIKKFFITTVLPAFG---GLNQLLWY 570
 QY 563 PEAPTPNATYESQVPEWGOBDESGNA-----QYRGLLEASLELFKFTTIGMELAFQE--- 614
 DB 571 -VAELEKKNICYHLHPVADFDDEKACTIWRFRSNLFTSOSLEFMAVSFGIADLVSPDLAG 629
 QY 615 -OLHFGWVLLLLLAYVLLTYILLMLALMSETVNSVATDSMTWKLOKA 665
 DB 630 IKSTFRMALLMGYSYVNIITVLLMLLMAIMSNSYQIISERADTEWKPARS 681

RESULT 3
 TRPL_DROME STANDARD; PRT; 1124 AA.
 AC P48994;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 30-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE TRANSIENT-RECEPTOR-POTENTIAL LIKE PROTEIN.
 GN TRPL.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON-R;

RX MEDLINE-92232293; PubMed-1314616;
 RA Phillips A.M., Bull A.L., Kelly L.E.;
 RT "Identification of a Drosophila gene encoding a calmodulin-binding
 protein with homology to the trp phototransduction gene";
 RL Neuron 8:631-642(1992).
 CC -1- FUNCTION: BINDS CALMODULIN. SUGGESTED TO MEDIATE CALCIUM ENTRY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE RHABDOMERIC
 CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 ANK REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: M8185; AAA28979.1; -
 CC FLYbase; FBgn0005614; trpl.
 CC InterPro: IPR002110; -
 CC InterPro: IPR002153; -
 CC Pfam: PF00023; ank; 2.
 CC PRINTS: PR01097; TRANSRECEPT.
 CC PROSITE: PS50088; ANK_REPEAT; 1.
 CC PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 CC KW Ionic channel; Transmembrane; Ion transport; Calcium channel;
 CC KW Calmodulin-binding; Vision; Ank repeat; Repeat.
 CC FT TRANSMEM 341 361 POTENTIAL.
 CC FT TRANSMEM 374 394 POTENTIAL.
 CC FT TRANSMEM 432 452 POTENTIAL.
 CC FT TRANSMEM 513 533 POTENTIAL.
 CC FT TRANSMEM 549 569 POTENTIAL.
 CC FT TRANSMEM 609 629 POTENTIAL.
 CC FT TRANSMEM 646 666 POTENTIAL.
 CC FT REPEAT 78 107 ANK 1.
 CC FT REPEAT 152 181 ANK 2.
 CC FT DOMAIN 710 727 CALMODULIN-BINDING (POTENTIAL).
 CC FT DOMAIN 809 825 CALMODULIN-BINDING (POTENTIAL).
 CC SQ SEQUENCE 1124 AA; 127697 MW; E14796DS5A2C10BD CRC64;

Query Match 4.9%; Score 197.5; DB 1; Length 1124;
 Best Local Similarity 19.7%; Pred. No. 2.2e-06;
 Matches 162; Conservative 130; Mismatches 281; Indels 249; Gaps 38;

QY 131 VNACILPQLQIDRDSGNPPL-----VNACSTD 159
 DB 25 VGCCVPL-----GLPQLLEKKFKLLAVERGMPNVRILQAKALRHQHNINCMDB 77
 QY 160 YRGHSALHAIERKSLQCVKLVENG---ANVHARAC-----GRFFQK 201
 DB 78 L--GRALFLAIDNENLEWVELLVVGVETKQALLHAINAEFEVALLHEBELYKKG 135
 QY 202 Q-----GTCFYGEL-PLSLAACKQMDVSYLLEN-----PH-----QPSL 238
 DB 136 EPYSMQKVDINTAFAPDITPLMLAAHKNFFELIRLLDRGAAPVPYPHDIRGCECEVRL 195
 QY 239 QATDSQNTVHL-----ALVMSIDNSAEINAVLSMTDGLQACARLCPTVQLD 288
 DB 196 TAEDSLRHSISRNVIRYALCSFSLICTSNDSSSTAFOLSWELRNALTEQDEKSKYMDL 255
 QY 289 IRNLQ---DLPLPKLAKEGKIEI-----FRHILQREFSGLSHLSKRTKFTWC 332
 DB 256 RQCCQKFAVNDLIDQRTSNELAILWYDPQMSYERGDMSLRLVQALSYKQKTV--A 313
 QY 333 YGFVRYSL---YDLASVDSCEENVLEIIFHCKSPHRHNVLEPLNKLK-----QAK 383
 DB 314 HSNIDQLSSIMYD--GLPGFRKRSIVDKVI--CIA-----QVAVLEPLCYLIYMCAPNCR 365

OY	384	MDLLIPPEFLFCLN----	LIVVEITFAVA-----	THQPLTKROAPHLKAEVGN	422
Dd	366	TGOLMRPKPMFKFLHASSYLFEFILILYSQRADDFVIRFGTTAKRKLAAQELRQNG			425
OY	430	SMLLTGHLILLGGIYLVVLQWLWFMKRHRVFIWISFIDSYEFIEFLFOALL--TVVSQVL			487
Dd	426	----TPSKLELIIVMVY-----IGFWVEVOEIPAVGMKSRYLRNMNNFIDFLNSLYSVM			477
OY	488	CFLAIEM-----YLP-----	LVLSALVLGMLNLTYTRBQG	518	
Dd	478	CLRPFATIQOATEIARDPOMAYVIPREKHMFDPOLIABLPAANVFSAKLKVHLFSINP			537
OY	519	HTGIYSVAIOKVIILRDRLRIILIYIFLGFCAVALISLSQEW-----REAPTGPNAAT			572
Dd	538	HILGFLQSLSGRMVY-DLVKFFFIYTLVLFAPA---CGLMQLMYFALESKSCYLLPG--			591
OY	573	ESVOPMGQEDEGNGA-----QYRGILEASILELKFPTIGCEL-----AFQOLHFPR			619
Dd	592	-----GEADMGSHGDSCMKMRFRGLFESSQSILFPWASFAGWGDDPELGSIKSYTRFW			644
OY	620	GMYLLLLAYLVLIYLIILLMLIALMSEFNVSATDSWISIMLKOAISYLEMENGYMCR			679
Dd	645	G-ILMGCSYVINIVYLLMLLIAAMSNSYAMIDEHSDETCKFAR-----TKLMASY			694
OY	680	KKORAGVLTGTGKPDGSDPPERMCFRV-----EEVNWASNEQTLLPIC			722
Dd	695	FEDSA-----TLPPFENVLPVKMWIRIFRKSSKTIDRSKKRRKBQEGSEVDNIMRSLV			750
OY	723	EDPSGAGVPTLENPVLASPKKDE--DGASENYVPVOLQ			762
Dd	751	W-RYVAAMHKRFEN---NFVEDDLINEVKSSETNTMYEMLE			787
RESULT	4				
TRPL_MOUSE	ID	TRPL_MOUSE	STANDARD;	PRT;	809 AA.
AC	061056;	035722;			
DT	01-OCT-2000	(Rel. 40,	Created)		
DT	01-OCT-2000	(Rel. 40,	Last sequence update)		
DT	01-OCT-2000	(Rel. 40,	Last annotation update)		
DE	TRANSIENT RECEPTOR POTENTIAL CHANNEL 1 (TRANSIENT RECEPTOR PROTEIN 1)				
DE	(MRP1) (TRP-RELATED PROTEIN 1).				
GN	TRPC1 OR TRRP1 OR TRPI.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.			
RC	TISSUE=Insulinoma;				
RA	MEDLINE=97307994;	PubMed=9155220;			
RA	Sakura H., Ashcroft F.M.;				
RT	"Identification of four trpl gene variants murine pancreatic beta-				
RT	cells.";				
RL	Diabetologia 40:528-532(1997).				
RN	[2]	SEQUENCE FROM N.A. (BETA ISOFORM).			
RC	TISSUE=Lens epithelium.				
RA	Rae J.L.;				
RT	"Ion channels in lens epithelia."				
RL	submitted (OCT-1999)				
RN	[3]	SEQUENCE OF 551-615 FROM N.A.			
RP	MEDLINE=96234226;	PubMed=8646775;			
RA	Zhu X., Jiang M., Peyton M., Boulay G., Hurst R., Stefan E.,				
RA	Birbaumer L.;				
RT	"Trp, a novel mammalian gene family essential for agonist-activated				
RT	capacitative Ca ²⁺ entry.";				
CC	Cell 85:661-671(1996).				
CC	-I- FUNCTION: SUGGESTED TO MEDiate CAPACITATIVE CALCIUM ENTRY (CCE).				
CC	SEEMS TO FORM A CALCIUM PERMEANT CHANNEL.				
CC	-I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).				
CC	-I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA (SHOWN HERE) AND BETA; ARE				

CC PRODUCED BY ALTERNATIVE SPLICING.
CC -I- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY.
CC -I- SIMILARITY: CONTAINS 3 ANR REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
CC or send an email to license@isb-sib.ch).

[illegible][illegible]

OY	256	SD-----	-----NSAENIALVTSMYGDLQAGARLPTVOLIEDIR-----	291
DB	242	ADLKELSLVEVEFRNDYELARQKMFANDLLAQAAR--NSRELFLVLTNTSSDEPLDKRG	299	
OY	292	----LQDITPLKLAKEGKIETFRHILQREESGSLHSKRKFTWCYGVAVSLYDLASVD	347	
DB	300	ILEERNINSLRLKLAIRYN-----OKEF-----	325	
OY	348	SEENSVLEIIIFHCKSPHRNR-----NVVLE-----PLNKL-----QAQMDLLIPKF	391	
DB	326	NCOQ--FLNTVWFGQSGYRKRPCTCKKITVTVLVTGFMFVLSLCYLIAKSGOGRITHP	383	
OY	392	FLNPLCNLTLYMEIFT-----AVAYHQPFLKQAAPHLKAEVGNMILGHILLILGIGIY	445	
DB	384	FMKFTIHGASGYFTPLLLNLNLSLYVNEDK-----KTMGPALERIDYLLIL-----	429	
OY	446	LLVGDLMTFWRHRYETWISFIDSYE-----ILFLFOA--LLTVVSQVLCFLAI-----	492	
DB	430	WITGIMWSDIKR--LWEGLEDFLEESRNOISFEVANSLSYLATFALKVVAHNKPHDFADR	486	
OY	493	EW--YLPLFVALVLGWMLNLLYTRGFQHTGIVSYM-----IQKYLIDBLRFLLLIYV	544	
DB	487	KQMDAFHPLTLVAGELGFAPRANVLSYLRLPFMYTSSILGLOISMGMLODPFGKFLGMFL	546	
OY	545	FLFGFVALVLSLSQEARPEALPTGPNAEVSYPOMEGQDEGNGAORYGILLESLELKEFT	604	
DB	547	VLFSEFTIGLTQYLDKGY-----TSKEQKDCVGFCEQGSNDT---FHSFTGCFALFWYI	598	
OY	605	IGMGELAF-----OEOLHFRGVVLLLLAYVLTITLLNMLTALMSERVNSVATQS	656	
DB	599	PSLAHVAFIVTFRTSGEELQSFVGA--IVGYTVVAVVLTKLIVAMLHKSPOLIATHE	656	
OY	657	WSIMKLOKA	665	
DB	657	DKEMKFARA	665	
RESULT	7			
AC	ANK1_MOUSE	STANDARD:	PRT: 1862 AA.	
AC	002357:			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	ANKYRIN 1 (ERYTHROCYTE ANKYRIN).			
GN	ANK1 OR ANK-1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NCBI	NCBI_Taxid=10090;			
RA	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Erythrocyte;			
RA	MEDLINE=92345717; Pubmed=1386265;			
RA	White R.A., Birkenmeier C.S., Peters L.L., Barker J.E., Lux S.E.;			
RT	"Murine erythrocyte ankyrin cDNA: highly conserved regions of the			
RT	regulatory domain."			
RL	Mamm. Genome 3:281-285(1992).			
CC	-I- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL			
CC	ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO			
CC	NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE			
CC	CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.			
CC	ERYTHROCYTE ANKIRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE			
CC	CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;			
CC	THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.			
CC	-I- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTE			
CC	PLASMA MEMBRANE.			
CC	-I- PUT: REGULATED BY PHOSPHORYLATION (BY SIMILARITY).			
CC	-I- PUT: ACYLATED BY PALMITIC ACID GROUP(S) (BY SIMILARITY).			
CC	-I- SIMILARITY: CONTAINS 23 ANK REPEATS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			

[illegible]

OY 377 ----NKLQAKMDLIPKFLNLIYMFETAVAYOPTLKOAPHLKA-----EY 427
 DB 572 AVHNNLDIVK--LLPRL-----GGSPHSPAMNGTPLHIAKQNOIEV 613
 OY 428 GNSMLTIG 435
 DB 614 ARSLLOYG 621
 RESULT 8
 TRP3_HUMAN STANDARD: PRT: 848 AA.
 ID TRP3_HUMAN 013507: 000593: Q15660;
 AC Q13507: 000593: Q15660;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DE 01-OCT-2000 (Rel. 40, Last annotation update)
 DE TRANSIENT RECEPTOR POTENTIAL CHANNEL 3 (TRP3-3).
 GN TRP3 OR TRP3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-9634226; PubMed-8646775;
 RA Zhu X., Jiang M., Peyton M., Boulay G., Hurst R., Stefani E.,
 RA Birbaumer L.;
 RT "Trp, a novel mammalian gene family essential for agonist-activated
 RT Cell 85:661-671(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97358541; PubMed-9215637;
 RA Xu X.-Z.S., Li H.-S., Guggino W.B., Montell C.;
 RT "Coassembly of TRP and TRPL produces a distinct store-operated
 RT conductance." J. Cell Biol. 115:1164(1997).
 RN [3]
 RP SEQUENCE OF 632-747 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE-96003837; PubMed-7568191;
 RA Wes P.D., Chevesich J., Jeromin A., Rosenberg C., Stetten G.,
 RA Montell C.;
 RT "TRP1, a human homolog of a Drosophila store-operated channel";
 RT Proc. Natl. Acad. Sci. U.S.A. 92:9652-9656(1995).
 CC -!- FUNCTION: SUGGESTED TO MEDIATE CAPACITATIVE CALCIUM ENTRY (CCE).
 CC -!- SEEMS TO FORM A CALCIUM PERMEANT CHANNEL.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY.
 CC -!- SIMILARITY: CONTAINS 4 ANK REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U47050; AAC51653.1;
 DR EMBL: Y13758; CA674083.1;
 DR EMBL: X89068; CAA61448.1;
 DR MIM: 602345;
 DR InterPro: IPR002110;
 DR InterPro: IPR002111;
 DR InterPro: IPR002153;
 DR Pfam: PF00023; ank; 2.
 DR PRINTS: PRO1097; TRANSRECEPTR.
 KW Ionic channel; Transmembrane; Ion transport; Calcium channel;
 KW ANK repeat; Repeat; Glycoprotein.
 FT TRANSMEM 350 370 POTENTIAL.
 FT TRANSMEM 382 402 POTENTIAL.

FT TRANSMEM 431 451 POTENTIAL.
 FT TRANSMEM 464 484 POTENTIAL.
 FT TRANSMEM 536 556 POTENTIAL.
 FT TRANSMEM 580 600 POTENTIAL.
 FT TRANSMEM 616 636 POTENTIAL.
 FT TRANSMEM 650 670 POTENTIAL.
 FT REPEAT 38 67 ANK 1.
 FT REPEAT 73 102 ANK 2.
 FT REPEAT 104 130 ANK 3.
 FT REPEAT 159 188 ANK 4.
 FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 416 416 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 739 742 EMGM -> GNGEM (IN REF. 3).
 SQ SEQUENCE 848 AA: 97354 MW: 1D8C9ZBC941DE416 CRC64;
 Query Match 3.5%; Score 139; DB 1; Length 848;
 Best Local Similarity 18.4%; Pred. No. 0.024;
 Matches 128; Conservative 96; Mismatches 222; Indels 248; Gaps 32;
 OY 152 VNAOCTDYYRGHSALHAIERSLQCYKLV--ENGANV-----HAR----- 192
 DB 65 LNVNCVD--YMGONMLQLVGNHLEVTLLKKEMLARIGALLAISKGVRIVEAIL 122
 OY 193 -----ACGRFQKG-----OGTCFPGELPLSLACFTQKQMVVSYLL-- 229
 DB 123 NHPFAASKRLTLPSCDELDDDFAYVDEGTRESPTITILLAHCKQYEVVHMLMK 182
 OY 230 ----ENPH-----OPASLOANDSQGNVLAHV 253
 DB 183 GARIEPHDYFCCKGDCMKORHDSFSRSHINAYKGLASAVAYSL--SSDPVLTALF 240
 OY 254 MISDMSAENIALVTSMYD-----GLLOAGARLC-PTVQLEDIRN--LQDL 295
 DB 241 L--SNELAKLANIEKEFKNDYRKLSMOCKDFVGVYLD--LCRDEEVEALINDLESA 294
 OY 296 PPLKLAAGKIEIFRH--ILOREFSGLSLSRKPFEM--CYGPVSLYDLASDSCSE 351
 DB 295 EPL-----EVHRKASLSRVKLAIRYEVKRYVAHPNCOQLTIWYE--NLSGURE 343
 OY 352 NSVLEIIAFHCKSPHRHRVVL-----EPLNKLQAKMDLIPKFL 393
 DB 344 ----QTIAIKC-----LVVLVALGUPFLAIGYWIAPCSLGLILRSP-----FM 384
 OY 394 NFLCLLIYFITTAVAYHO-----PYLKQOAPHLKAEVGNLSMLTGHLIL-- 440
 DB 385 KEVAHAASEPIIFGLLVFNASDRFEGITTLPLNITVTDYPRKQIFRYKTOFTWEMLIWV 444
 OY 441 -LGGIYLVGOLWY-----FMRHFEI-----WSF-----IDSY 469
 DB 445 VLGKMSSECKELMECPREYIQLNVLDPFGLSLFIAFAFARLAFLOAKQOQYVDSY 504
 OY 470 FFLFLFOALLLVVSOVLCFLAIEWLP-----LLVSALVGLMNLVLYTGTGFOHT 520
 DB 505 VQESDLSVTLPELQYFTYARDKM-LPSDDPQIIEGLATVAIVLSFSAIATILANESF 563
 OY 521 GIYSWIKQVILRDLRLFLLYLVFLGFAVALVLSOEAMPEAPTPGNATESVOAMEG 580
 DB 564 GPLQISLGRV-KDIFKFMVLEIMVFAPFMIGMFTLYSYLGAKV----- 607
 OY 581 QDEDENGAQYKQILASLELFFETIGMELA-----FOELHFRGMVLLALLAVL 631
 DB 608 -----NAFTVSESEFTLSEVSIPLSEVSVLVKDYHKFTENIGY-----VLGIYIV 656
 OY 632 LTYILLNMLALMSETVNSVATDSMSIMKLOKA 665
 DB 657 TMVVVLLNMLIAMINSYOEIEDSDVEMKPARS 690
 RESULT 9

ANK1_HUMAN STANDARD: PRT: 1880 AA.
 ID ANK1_HUMAN
 AC P16157;
 DT 01-APR-1990 (rel. 14, Created)
 DT 01-APR-1990 (rel. 14, Last sequence update)
 DT 01-OCT-2000 (rel. 40, Last annotation update)
 DE ANKYRIN 1 (ERYTHROCYTE ANKYRIN) (ANKYRIN R) (ANKYRINS 2.1 AND 2.2).
 GN ANK1 OR ANK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.
 RC TISSUE=Hematopoietic;
 RX MEDLINE=90158830; PubMed=2137557;
 RA Lux S.E., John K.M., Bennett V.;
 RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
 RT structure with homology to tissue-differentiation and cell-cycle
 RT control proteins.";
 RL Nature 344:36-42(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=90175370; PubMed=1689849;
 RX Lambert S., Yu H., Prchal J.T., Lawler J., Ruff P., Speicher D.,
 RA Cheung M.C., Kan Y.W., Palek J.;
 RT "cDNA sequence for human erythrocyte ankyrin";
 RT Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).
 RN [3]
 RP VARIANT HS IL6-462.
 RX MEDLINE=96225450; PubMed=8640229;
 RA Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,
 RA Dornwell M., Herbers J., Kugler W., Oezcan R., Pekrun A.,
 RA Gallagher P.G., Schroeter W., Forget B.G., Lux S.E.;
 RT "Ankyrin-1 mutations are a major cause of dominant and recessive
 RT hereditary spherocytosis.";
 RL Nat. Genet. 13:214-218(1996);
 CC -1- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
 CC ELEMENTS: BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO
 CC NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN G85, AND TO THE
 CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
 CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
 CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
 CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
 CC PLASMA MEMBRANE.
 CC -1- ALTERNATIVE PRODUCTS: VARIOUS ISOFORMS OF ANKYRIN ARE PRODUCED BY
 CC ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF ANKYRIN
 CC VARIANT 2.1.
 CC -1- PTM: REGULATED BY PHOSPHORYLATION.
 CC -1- PTM: ACYLATED BY PALMITIC ACID GROUP(S).
 CC -1- DISEASE: DEFECTS IN ANK1 ARE THE CAUSE OF DOMINANT AND RECESSIVE
 CC HEREDITARY SPHEROCYTOSIS (HS).
 CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X16609; CAA34610.1; -
 DR EMBL: M28880; AAA51732.1; -
 DR PIR: S08275; SJHUK.
 DR PIR: A35049; A35049.
 DR HSSP: 000420; 1AWC.
 DR MIM: 182900; -
 DR InterPro: IPR000488; -
 DR InterPro: IPR000906; -
 DR InterPro: IPR002110; -
 DR Pfam: PF00791; Z05; 1.

DR Pfam: PF00023; ank: 22.
 DR Pfam: PF00531; death: 1.
 DR PROSITE: PS50088; ANK_REPEAT: 20.
 DR PROSITE: PS50297; ANK_REPEAT_REGION: 1.
 DR PROSITE: PS50017; DEATH_DOMAIN: 1.
 DR Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
 KW Phosphorylation; Lipoprotein; Multigene family; Disease mutation;
 KM Elliptocytosis; polymorphism.
 FT INIT_MER 0
 FT DOMAIN 1 826
 FT
 FT DOMAIN 827 1381
 FT
 FT DOMAIN 1382 1880
 FT
 FT REPEAT 43 72
 FT REPEAT 76 105
 FT REPEAT 109 138
 FT REPEAT 142 171
 FT REPEAT 173 200
 FT REPEAT 204 233
 FT REPEAT 237 266
 FT REPEAT 270 299
 FT REPEAT 303 332
 FT REPEAT 336 365
 FT REPEAT 369 398
 FT REPEAT 402 431
 FT REPEAT 435 464
 FT REPEAT 468 497
 FT REPEAT 501 530
 FT REPEAT 534 563
 FT REPEAT 567 596
 FT REPEAT 600 629
 FT REPEAT 633 662
 FT REPEAT 666 695
 FT REPEAT 699 728
 FT REPEAT 732 761
 FT REPEAT 765 794
 FT DOMAIN 1402 1486
 FT VARSPLIC 1512 1873
 FT VARSPLIC 1874 1874
 FT VARSPLIC 1849 1880
 FT
 FT VARIANT 20 20
 FT
 FT VARIANT 462 462
 FT
 FT VARIANT 618 618
 FT
 FT VARIANT 749 749
 FT
 FT VARIANT 844 844
 FT
 FT VARIANT 1391 1391
 FT
 FT VARIANT 1285 1285
 FT
 FT VARIANT 1591 1591
 FT
 FT VARIANT 1698 1698
 FT
 FT CONFLICT 229 229
 FT CONFLICT 1345 1345
 SQ SEQUENCE 1880 AA; 206145 MW; 1C5F5E7BED1CD428 CRC64;

Query Match 3.4%; Score 138; DB 1; Length 1880;
 Best Local Similarity 22.8%; Pred. No. 0.079; Indels 110; Gaps 16;
 Matches 82; Conservative 45; Mismatches 122;
 QY 134 CILPLQIDRD-----SGNDFPL-VNAOCTDYYR-----GHSALHI 169

```

Db 319 CVALLLQYAEIDITPLDHTPLHVAHC--GHHRAKVLDDGAKPNSRALGFTPLHI 376
OY 170 ALEKRSLOCVKLVENGANY-----HARAGREFPOK--GGTCFYFG 209
Db 377 ACKKNHVRWELLKTKGASIDAVTESGLPLPHVASFMGHLPIYKNLLDORASPNVSNKY 436
OY 210 ELPLSLACTKQMDVSYLLENPHOPASLOATDSQNTVLAHALMTSDNSAENIALVTSW 269
Db 437 EFTPLHMAARGHTEVAKYLLDN---KAKVNAKAKKDDQTPHCAARGH-----TNK 484
OY 270 YGCLLAGARLCTVQLEDIRNLODITPLKLAKEGKIEIFRITLQREFSGLSLSRKPT 339
Db 485 VKLLENNMN-----PMLATTAGHTPLHIAAREGHVETVIALLEKEASQACMTKKGFT 537
OY 330 EW-----CYGPRVRS---LYDLASVDSCEENSVLEI--IAFHCKSPHHRVWVLEPLKLLQ 381
Db 538 PLHVAARYKGVRAVELLERDAHPNAGKNGLPLHVAH-----HNNLDIV----- 564
OY 382 AKWDLIPREFLNLCLYMFIFTAVAAYHOPTLKQAAPHLKA-----EVGNSMLLTG 435
Db 585 ---KILLPR-----GGSPHSPAMNGYTPHLHAKQNOYEVARSLLQYG 624

RESULT 10
LATA_LATMA STANDARD; PRT; 1401 AA.
ID LATA_LATMA
AC P23631;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ALPHA-LATROTOXIN PRECURSOR.
OS Latrodectus mactans (Black widow spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
  Araneomorphae; Entelegynae; Araneidae; Theridiidae; Latrodectus.
OX NCBI_TaxID=6924;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TREDECINGUTATUS; TISSUE=Venom;
RX MEDLINE=91031994; PubMed=1977615;
RA Kiyatkin N.I., Dulubova I.E., Chekhovskaya I.A., Gishin E.V.;
RT "Cloning and structure of cDNA encoding alpha-latrotoxin from black
  widow spider venom."
  FEBS Lett. 270:127-131(1990).
  [2]
RN [2]
RP PARTIAL SEQUENCE.
RX MEDLINE=91362695; PubMed=1888339;
RA Volkova T.M., Galikina T.G., Kudelin A.B., Nazimov I.V., Gishin E.V.;
RT "Structure of tryptic fragments of a neurotoxin from black widow
  spider venom."
  Bioorg. Khim. 17:437-441(1991).
CC -1- FUNCTION: CAUSES SYNAPTIC VESICLES EXOCYTOSIS AND
  NEUROTRANSMITTER RELEASE FROM PRESYNAPTIC NERVE TERMINALS.
CC -1- PTM: ALPHA-LATROTOXIN MAY BE PROCESSED AT ITS C-TERMINUS.
CC -1- SIMILARITY: CONTAINS 21 ANK REPEATS.
DR HSSP; 000420; IAMC.
DR InterPro; IPR002110; -.
DR Pfam; PF00023; ank. 17.
DR PROSITE; PS50088; ANK_REPEAT; 11.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KM Toxin; Neurotoxin; Repeat; ANK repeat; Venom; Signal.
FT SIGNAL 1 20
FT CHAIN 21 1401 ALPHA-LATROTOXIN.
FT REPEAT 490 521 ANK 1.
FT REPEAT 525 554 ANK 2.
FT REPEAT 559 589 ANK 3.
FT REPEAT 593 622 ANK 4.
FT REPEAT 626 656 ANK 5.
FT REPEAT 660 690 ANK 6.
FT REPEAT 695 723 ANK 7.
FT REPEAT 729 758 ANK 8.
FT REPEAT 762 791 ANK 9.

```

```

FT REPEAT 795 824 ANK 10.
FT REPEAT 828 857 ANK 11.
FT REPEAT 862 891 ANK 12.
FT REPEAT 895 924 ANK 13.
FT REPEAT 928 957 ANK 14.
FT REPEAT 971 1003 ANK 15.
FT REPEAT 1003 1033 ANK 16.
FT REPEAT 1035 1064 ANK 17.
FT REPEAT 1068 1097 ANK 18.
FT REPEAT 1101 1131 ANK 19.
FT REPEAT 1137 1166 ANK 20.
FT REPEAT 1170 1199 ANK 21.
SQ SEQUENCE 1401 AA; 156834 MW; 21CE198B80172568 CRC64;

Query Match
Best Local Similarity 3.4%; Score 134.5; DB 1; Length 1401;
Matches 69; Conservative 62; Mismatches 123; Indels 89; Gaps 13;

OY 35 GLPPMESOPGEGDR--KAPQIRVNLNRYRGT-----GASQPD 70
Db 594 GTFPLHYAIRGGERILEAFNLQSIDVNAKSNMGLPFLHAIKKNDWPVASTLLGSKVD 653
OY 71 PNRFDRLFNVAVSRGVPEDLAQLPEYLSKTSKYLTDSSETGSGTKLMK---AVLN 126
Db 654 INAVDENNI-----TALHYAAILGYLETFKOLINLKEINAVVSSPELLSALHYAIIY 706
OY 127 LKDVAVACILPLLOIDRDSGNQPLVYNACSTDYHGHSAHLHAIKRSLOCVKLVENG 186
Db 707 KHDVASFLMRSSNV-----VNLKALG---GTFPLHVAIYQGRKQILSLMFQDG 753
OY 187 ANVHARACGRFPOKGGCTFYFGLPLSLACTKQMDVSYLLENPHOPASLOATDSQGN 246
Db 754 VINEOKTDEKY-----PFLHIAAMSKYPELIQIIDL---QGSNFEAKTNSGA 797
OY 247 TVLH-----ALVMSD-----NSAEN--IALVTSMDGLQ--AGARLCPYQLED 288
Db 798 TPLHLATFKGSOAALILINNENVMRDTEGCMPIHGAAMGLDVAQAIIISIDATVVD 857
OY 289 INNLDTPLKLAKEGKIEIFRHILQ-----REFSLSHL 324
Db 858 IDKNSDPLNLAQNSHDVIAKIFYDQAGADINTRNKKGLAPL 900

RESULT 11
YAZA_SCHPO STANDARD; PRT; 642 AA.
ID YAZA_SCHPO
AC Q09701;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 72.5 KDA PROTEIN C2F7.10 IN CHROMOSOME 1.
GN SPAC2F7.10.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
  Schizosaccharomycetales; Schizosaccharomycetaceae;
  OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RT Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO YEAST AKR1.
CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation
  at the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb.ch/announce/
  or send an email to license@isb.slb.ch).
  CC -----

```

```

CC -----
DR EMBL: 250142: CAA90497.1: -
DR HSSP: Q00420: IANC
DR InterPro: IPR001594: -
DR InterPro: IPR002110: -
DR Pfam: PF00023: ank; 6.
DR Pfam: PF01529: zf-DHHC; 1.
DR PROSITE: PS50088: ANK_REPEAT; 1.
DR PROSITE: PS50297: ANK_REPEAT_REGION; 1.
DR Hypothetical protein; ANK repeat; Repeat.
FT REPEAT 1 29 ANK 1.
FT REPEAT 33 62 ANK 2.
FT REPEAT 67 96 ANK 3.
FT REPEAT 100 129 ANK 4.
FT REPEAT 133 162 ANK 5.
FT REPEAT 166 196 ANK 6.
SQ SEQUENCE 642 AA: 72521 MW: C83584A3300BDDA0 CRC64;

```

```

Query Match 3.3% Score 133.5; DB 1: Length 642;
Best Local Similarity 21.0%; Pred. No. 0.042;
Matches 95; Conservative 71; Mismatches 136; Indels 151; Gaps 23;

```

```

QY 152 VNAQCTDDYRGSHALHIAIEKRSLOCVLLVNGANVHARACGRFGKGGTCFYFGEI 211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 25 IDVNATDE--GGATLHMAALNQIPICKFLEHGADVNAIG-----GDL 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 212 ---PLSLACTKQMDVSVLLENPHOPASIQATDSOGNTVLAHVAISD----- 257
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 68 QAAPHIMAKRSGSVKTVHLYOHGADPL--LKDQGFNCLHVAHVAAPLVAHLLHD 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 258 ---NSANIALVTSMDG-----LLOAGARLCPTVQLEDINLQDLPFLKLAKE 304
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 125 ISVDLRDDQHTPLMASHNGNPIINCLLRGADVLAD-----DKMPLHMSIYG 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 305 GKIEIFRHILQREFSGL-----SHLSRKF-TEWCY-GPVARS-LYDLASVDS--CEEN 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 178 GNLCKMKLILKE--GGICTAVANLSGQLKFPWALASELRVSHLRFKQALISNGKVKET 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 353 SVLEIIFAFCKSPHRRHMYVLEPLNKLQAKMDLLPKFELNLCMLIYFETTAAYRH 412
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 236 SE-----EP-----EKWVVSRSKQFSOKTPIICFL----- 262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 413 PTLKQAAPHLKAEGNSMLTGHILIG---GIYLLVGLM-YFWR----- 457
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 263 -----SFTITGVFFIMSICPMVSLIAPRMIIYFTKYYITTCIHANI 306
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 458 ---HVEFIWISFIDSYELLFLF---QALLTVVSQVLCFLAIEWYLDPLVASLVGLMLNL 511
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 307 DLYHFELETPFLAGISFIFFWWCHSLIYVKTLPKIPKLSLFLVLSFTCIG---L 362
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 512 YTRGFQHTG---IYSW-----IOKVLRDL 535
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 363 YVTAFAQNGVYDKIGAVVORREESIKLKDGL 395
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 12
TRP6_MOUSE STRAND: PRT; 930 AA.
AC Q61143: Q922J1:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TRANSIENT RECEPTOR POTENTIAL CHANNEL 6 (CALCIUM ENTRY CHANNEL).
GN TRP6 OR TRP6 OR TRP6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98037793; PubMed=9368034;

```

```

RA Boulay G., Zhu X., Peyton M., Jiang M., Hurst R., Stefani E.,
RA Birbaumer L.:
RT "Cloning and expression of a novel mammalian homolog of Drosophila
RT transient receptor potential (trp) involved in calcium entry secondary
RT to activation of receptors coupled by the Gq class of G protein.";
RL J. Biol. Chem. 272:29672-29680(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DRA/2;
RX MEDLINE=99158172; PubMed=10050885;
RA Buess M., Engler O., Hirsch H.H., Moroni C.;
RT "Search for oncogenic regulators in an autocrine tumor model using
RT differential display PCR: identification of novel candidate genes
RT including the calcium channel mtrp6.";
RL Oncogene 18:1487-1494(1999).
RN [3]
RP SEQUENCE OF 631-739 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96234226; PubMed=8646775;
RA Zhu X., Jiang M., Peyton M., Boulay G., Hurst R., Stefani E.,
RA Birbaumer L.:
RT "trp, a novel mammalian gene family essential for agonist-activated
RT capacitative Ca2+ entry.";
RL Cell 85:661-671(1996).
CC -I- FUNCTION: NONSELECTIVE CAPACITATIVE CALCIUM ENTRY CHANNEL SUBUNIT.
CC MEDIATES CALCIUM ENTRY STIMULATED BY A G-PROTEIN COUPLED RECEPTOR
CC BUT NOT BY INTRACELLULAR CALCIUM STORE DEPLETION. ACTIVATED BY
CC DIACYLGLYCEROL (DAG) IN A MEMBRANE-DELIMITED FASHION INDEPENDENTLY
CC OF PROTEIN KINASE C. IT IS PERMEABLE FOR CALCIUM, CESIUM, SODIUM,
CC POTASSIUM AND MAGNESIUM.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -I- TISSUE SPECIFICITY: LUNG AND BRAIN.
CC -I- PTM: N-GLYCOSYLATED.
CC -I- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY.
CC -I- SIMILARITY: CONTAINS 3 ANK REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U49069: AAC06146.1: -
DR EMBL: AF057748: AAC64394.1: -
DR MGD: MGI:109523: trp6.
DR InterPro: IPR002110: -
DR InterPro: IPR002153: -
DR Pfam: PF00023: ank; 2.
DR PRINTS: PR01097: TRNSRECEPTRP.
DR PROSITE: PS50088: ANK_REPEAT; 1.
DR PROSITE: PS50297: ANK_REPEAT_REGION; 1.
DR Ionic channel; transmembrane; Ion transport; Calcium channel;
KW ANK repeat; Repeat; Glycoprotein.
FT TRANSMEM 406 426 POTENTIAL.
FT TRANSMEM 438 458 POTENTIAL.
FT TRANSMEM 467 507 POTENTIAL.
FT TRANSMEM 521 541 POTENTIAL.
FT TRANSMEM 592 612 POTENTIAL.
FT TRANSMEM 636 656 POTENTIAL.
FT TRANSMEM 706 726 POTENTIAL.
FT REPEAT 131 160 ANK 1.
FT REPEAT 162 188 ANK 2.
FT REPEAT 217 246 ANK 3.
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 727 727 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 3 56 MISSING (IN REF. 2).
FT CONFLICT 105 105 V -> A (IN REF. 2).

```

FT CONFLICT 114 114 W -> R (IN REF. 2).
 FT CONFLICT 134 134 N -> D (IN REF. 2).
 FT CONFLICT 184 184 S -> A (IN REF. 2).
 FT CONFLICT 371 371 D -> Y (IN REF. 2).
 FT CONFLICT 436 437 PR -> RG (IN REF. 2).
 FT CONFLICT 905 905 S -> T (IN REF. 2).
 SQ SEQUENCE 930 AA: 106732 MW: CF21M426972732F3 CRC64;

Query Match 3.2% Score 130; DB 1; Length 930;
 Best Local Similarity 17.8%; Pred. No. 0.12;
 Matches 124; Conservative 98; Mismatches 220; Indels 254; Gaps 29;

QY 152 VNAOCTDDYRGSHAHIAIEKSLQCVKLV--ENGANY-----HARACGRF- 197
 DB 123 LWNVCND--YMGONALQAVANHEHETELLKKNLSVGCALLAISKGYRIWEAIL 180
 QY 198 ----FOKGO-----GTCEYFEGELPLSLAACKOMDVSYL-- 229
 DB 181 NHPSEAEGRKATSPSOSLQDDFYAYDEGDFRSHDVTPLLAHCOYEIVHTLKK 240
 QY 230 ----EMPH-----QPASLQATPSQGNVYHALV 253
 DB 241 GARIEKPHDYFCKCTECSKOKHDSHSRSTRINAYKGLASPAYISL--SSEDPVMTALE 298
 QY 254 MISD-----NSAENIAL--VTSMYDGLAQARLC--PTVQLEDIRN----- 291
 DB 299 LBNELAVLANIEKEFNDRKLSMOCKDFVGLD--LCRTEVEVALINDAETROP 354
 QY 292 ----LODLPLKLAKEGIEIFRH-----ILQREPSGLSHLSRFTWCYGPVAV 338
 DB 335 GDFGRNLSKALAIKDEYKKFVAPHNCOOLSIWEMISGLRQOTMAVKFLVLAVAL 414
 QY 339 SLYDASVDSCENSVLEITAFHCKSPHRHNVLEPLKLLQAKMDLIPKFLMFLCN 398
 DB 415 GLPEFLALIWCAPCS-----KMGKILPREFMKFAVH 445
 QY 399 LIYMEFTAVAYHOPTLKRQAAPHKAEVG-----NSMLTGHILLGLGCIYLL 447
 DB 446 AASFTEFLGLVWNNADREGTKLPNETSTONARQLFRMKTSCFSMMMLI----ISWV 501
 QY 448 VQQLV---YFM-----RRNVFTWISIDSYFELFLQALLVVSQVLCF----- 489
 DB 502 IGMVAEKEIKWTGPKREYLFELMWNLD--FGMLAFFA--SFIAEFMAFMHASKAOSII 557
 QY 490 -----LAIEW--YLPLVVS-----ATVLGLNLLYVRGFO 518
 DB 558 DANDPLKDLTKYTLGDNVYVYLARKWDPDPQIISGLYALAVLSFRIAYIILPANE 617
 QY 519 HTGIYSVMIQKVLRLDLLELLIYVLEFGFAVALVLSQEAHREAPTPGNATESVQPM 578
 DB 618 SFGPLQISIGRTV-KDIEFVYIFIMVEFAFMIGMFLNLSY-----IGAKONEAFTIV 670
 QY 579 EGQDEBGNAGYRGILNLSLEFKFTIGMEL-----AFQDQHRGAVLLILLAY 629
 DB 671 E-----ESKTLFMAIFGLSEVKSVAIVYNHKEFTENIGY-----VLGYV 710
 QY 630 VLLFYILLNMLALMSEFVNSVATDSWMSIKLQKA 665
 DB 711 NVTWIVIVLLMLMIMINSFOEIEDDADAVEMKFAFA 746

RESULT 13
 NTC4 MOUSE STANDARD: PRT: 1964 AA.
 AC P31695: 062389:
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 4 PRECURSOR (TRANSFORMING
 DE PROTEIN INT-3).
 GN NOTCH4 OR INT3 OR INT-3.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92194507; PubMed=1312643;
 RA Robbins J., Blondel B.J., Callahan D., Callahan R.;
 RT "Mouse mammary tumor gene int-3: a member of the notch gene family
 RT transforms mammary epithelial cells.";
 RL J. Virol. 66:2594-2599(1992).
 RN [2]
 RP REVISIONS, SEQUENCE FROM N.A.
 RX MEDLINE=97294599; PubMed=9150355;
 RA Callahan D., Callahan R.;
 RT "The mouse mammary tumor associated gene INT3 is a unique member of
 RT the NOTCH gene family (NOTCH4).";
 RL Oncogene 14:1883-1890(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung, and Testis;
 RX MEDLINE=96281668; PubMed=8681805;
 RA Uytendaele H., Marazzi G., Wu G., Yan Q., Sassoon D., Kitajewski J.;
 RT cell-specific mammalian Notch gene.";
 RT Development 122:2251-2259(1996).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- DISEASE: ACTIVATED INT-3 TRANSFORMS MAMMARY EPITHELIAL CELLS.
 CC -1- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
 CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M80456; AAB38377.1; -
 DR EMBL: U43691; AAC52630.1; -
 DR PIR: A38072; TWNT3.
 DR HSP: P00740; 11XA.
 DR MGD: MGI:107471; Notch4.
 DR InterPro: IPR000152; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR000800; -
 DR InterPro: IPR001438; -
 DR InterPro: IPR001811; -
 DR InterPro: IPR002110; -
 DR Pfam: PF00008; EGF; 27.
 DR Pfam: PF00023; ank; 6.
 DR Pfam: PF00066; notch; 2.
 DR PRINTS: PR00010; EGFBLD.
 DR PROSITE: PS50088; ANK_REPEAT; 5.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 11.
 DR PROSITE: PS00022; EGF_1; 28.
 DR PROSITE: PS01186; EGF_2; 21.
 DR PROSITE: PS01187; EGF_CA; 9.
 KW Differentiation; Neurogenesis; Repeat; EGF-like domain; Transmembrane;
 KW Glycoprotein; Proto-oncogene; ANK repeat; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 FT EGF-LIKE 6.

FT	DOMAIN	273	309	EGF-LIKE 7,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	311	350	EGF-LIKE 8,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	352	388	EGF-LIKE 9,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	389	427	EGF-LIKE 10,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	429	470	EGF-LIKE 11,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	472	508	EGF-LIKE 12,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	510	546	EGF-LIKE 13,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	548	584	EGF-LIKE 14,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	586	622	EGF-LIKE 15,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	622	656	EGF-LIKE 16,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	658	686	EGF-LIKE 17,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	688	724	EGF-LIKE 18,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	726	762	EGF-LIKE 19,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	764	800	EGF-LIKE 20,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	803	839	EGF-LIKE 21,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	841	877	EGF-LIKE 22,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	878	924	EGF-LIKE 23,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	926	962	EGF-LIKE 24,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	964	1000	EGF-LIKE 25,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	1002	1040	EGF-LIKE 26,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	1042	1081	EGF-LIKE 27,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	1083	1122	EGF-LIKE 28,	CALCIUM-BINDING	(POTENTIAL)
FT	DOMAIN	1126	1167	EGF-LIKE 29,	CALCIUM-BINDING	(POTENTIAL)
FT	REPEAT	1168	1208	LIN/NOTCH 1,	CALCIUM-BINDING	(POTENTIAL)
FT	REPEAT	1209	1242	LIN/NOTCH 2,	CALCIUM-BINDING	(POTENTIAL)
FT	REPEAT	1243	1282	LIN/NOTCH 3,	CALCIUM-BINDING	(POTENTIAL)
FT	REPEAT	1628	1657	ANK 1,	CALCIUM-BINDING	(POTENTIAL)
FT	REPEAT	1661	1691	ANK 2,	CALCIUM-BINDING	(POTENTIAL)
FT	REPEAT	1695	1724	ANK 3,	CALCIUM-BINDING	(POTENTIAL)
FT	REPEAT	1728	1757	ANK 4,	CALCIUM-BINDING	(POTENTIAL)
FT	REPEAT	1761	1790	ANK 5,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	25	38	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	32	48	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	50	59	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	65	77	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	71	100	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	102	111	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	119	130	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	124	140	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	142	151	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	157	168	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	162	177	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	179	188	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	195	208	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	202	217	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	219	228	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	235	246	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	240	259	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	261	270	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	277	288	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	282	297	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	299	308	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	315	329	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	323	338	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	340	349	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	356	367	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	361	376	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	378	387	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	393	404	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	398	415	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	417	426	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	433	449	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	443	458	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	460	469	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	476	487	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	481	496	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	498	507	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	514	525	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	519	534	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	536	545	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	552	563	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	557	572	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)
FT	DISULFID	574	583	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL)

Query	Similarity	Score	DB	Length
Best Local	72: Conservative	37: Mismatches	103: Indels	70: Gaps
Query Match	Similarity 25.5%	Pred. No. 0.37		
DB	1553 LPQAMLPTRPOCESEBVLVDVTCGPDGVTPLMSAVFCGGVOYOTGASPORGLD-----	145		
OY	94 LPEYSKTSKYLTDBEY----TEGSTGKCLMKAV----LNKDGYNACILPLQIDRDS	145		
OY	146 GNPQP---LVNAOCTDby--RGHSALHIAIEKRSLOQVKLLVNGAN-----	188		
DB	1606 GNLEPMEPLLDGACPOAHVTCTGEPHLHARFSRPTARLLLEAGNPQDRAGRTP	1665		
OY	189 -----VHARACGRFPQGGQCTCF----YFGLPPLSLACTQMDVSYLLENPHQPSL	238		
DB	1666 LHTAVADAREVCOILLASRQTSVDARTEDGTPPLMLARLAVEDLVEPLA---ARADV	1722		
OY	239 QATDSOGNVTALHAIWMISDMSAENIALVTSWYDGLLOAGARCLPVTOLIEDINLODLPPL	298		
DB	1723 GARDKRGKTALHMAAANVNNARAAR-----SLLQAGAD-----KDAQDSREQTPPL	1766		
OY	299 FLAAREGKIEIFRHILQ-----REFSGLS---HLSRKFTW	331		
DB	1767 FLAAREGAVEVAQILLLEIGARGRLDQAGLADGAVARQSHW	1808		


```

RT "Restriction of P-element insertions at the Notch locus of Drosophila
RL melanogaster."
RN Mol. Cell. Biol. 7:1545-1548(1987).
RP [5]
RA REVIEW.
RT Harris W.A.;
RL "Many cell types specified by Notch function."
RT Curr. Biol. 1:120-122(1991).
CC -1- FUNCTION: NOTCH PROTEIN IS ESSENTIAL FOR PROPER DIFFERENTIATION OF
CC ECTODERM.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO
CC THE INNER PART OF EMBRIO IS ONE OF THE FIRST STEPS OF CNS
CC DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE
CC NEUROGENIC GENES.
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M16152; AAB59220.1; -.
DR EMBL: M16153; AAB59220.1; JOINED.
DR EMBL: M16149; AAB59220.1; JOINED.
DR EMBL: M16150; AAB59220.1; JOINED.
DR EMBL: M16151; AAB59220.1; JOINED.
DR EMBL: M16151; AAB59220.1; JOINED.
DR EMBL: M13689; AAA28725.1; -.
DR EMBL: M13689; AAA28725.1; JOINED.
DR EMBL: M13689; AAA28725.1; JOINED.
DR EMBL: M12175; AAA74496.1; -.
DR EMBL: M16025; AAA28726.1; -.
DR PIR: A24420; A24420.
DR PIR: A24768; A24768.
DR PIR: A05267; A05267.
DR HSSP: P00740; 11XA.
DR Flybase: FBgn004647; N.
DR InterPro: IPR000152; -.
DR InterPro: IPR000561; -.
DR InterPro: IPR000800; -.
DR InterPro: IPR001438; -.
DR InterPro: IPR001881; -.
DR InterPro: IPR002110; -.
DR Pfam: PF00008; EGF; 36.
DR Pfam: PF00023; ank; 6.
DR Pfam: PF00066; notch; 3.
DR PRINTS: PR00010; EGFBLD.
DR PROSITE: PS30088; ANK_REPEAT; 5.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 22.
DR PROSITE: PS00022; EGF_1; 34.
DR PROSITE: PS01186; EGF_2; 28.
DR PROSITE: PS01187; EGF_CA; 22.
KW Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Signal; Glycoprotein.
KW SIGNAL 1 44
FT CHAIN 45 2703 POTENTIAL.
FT DOMAIN 45 1745 NEUROGENIC LOCUS NOTCH PROTEIN.
FT TRANSMEM 1746 1765 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1767 2703 POTENTIAL.
FT DOMAIN 58 95 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 96 136 EGF-LIKE 1.
FT DOMAIN 139 176 EGF-LIKE 2.
FT DOMAIN 177 215 EGF-LIKE 3.
FT DOMAIN 217 253 EGF-LIKE 4.
FT DOMAIN 255 291 EGF-LIKE 5.
FT DOMAIN 293 329 EGF-LIKE 6.
FT DOMAIN 293 329 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).

```

```

FT DOMAIN 331 370 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 372 408 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 409 447 EGF-LIKE 10.
FT DOMAIN 449 486 EGF-LIKE 11.
FT DOMAIN 488 524 EGF-LIKE 12.
FT DOMAIN 526 562 EGF-LIKE 13.
FT DOMAIN 564 600 EGF-LIKE 14.
FT DOMAIN 602 637 EGF-LIKE 15.
FT DOMAIN 639 675 EGF-LIKE 16.
FT DOMAIN 677 713 EGF-LIKE 17.
FT DOMAIN 715 751 EGF-LIKE 18.
FT DOMAIN 753 789 EGF-LIKE 19.
FT DOMAIN 791 827 EGF-LIKE 20.
FT DOMAIN 829 865 EGF-LIKE 21.
FT DOMAIN 867 905 EGF-LIKE 22.
FT DOMAIN 907 944 EGF-LIKE 23.
FT DOMAIN 946 982 EGF-LIKE 24.
FT DOMAIN 984 1020 EGF-LIKE 25.
FT DOMAIN 1022 1058 EGF-LIKE 26.
FT DOMAIN 1060 1096 EGF-LIKE 27.
FT DOMAIN 1098 1134 EGF-LIKE 28.
FT DOMAIN 1136 1181 EGF-LIKE 29.
FT DOMAIN 1183 1219 EGF-LIKE 30.
FT DOMAIN 1221 1257 EGF-LIKE 31.
FT DOMAIN 1259 1295 EGF-LIKE 32.
FT DOMAIN 1297 1335 EGF-LIKE 33.
FT DOMAIN 1337 1373 EGF-LIKE 34.
FT DOMAIN 1375 1412 EGF-LIKE 35.
FT DOMAIN 1415 1451 EGF-LIKE 36.
FT REPEAT 1475 1513 LIN/NOTCH 1.
FT REPEAT 1514 1553 LIN/NOTCH 2.
FT REPEAT 1554 1593 LIN/NOTCH 3.
FT REPEAT 1601 1645 ANK 1.
FT REPEAT 1646 1690 ANK 2.
FT REPEAT 1691 1735 ANK 3.
FT REPEAT 1736 1780 ANK 4.
FT REPEAT 1781 1825 ANK 5.
FT REPEAT 1826 1870 ANK 6.
FT REPEAT 1871 1915 POLY-GLN (OPR-REPEAT).
FT REPEAT 1916 1960 BY SIMILARITY.
FT REPEAT 1961 2005 BY SIMILARITY.
FT REPEAT 2006 2050 BY SIMILARITY.
FT REPEAT 2051 2095 BY SIMILARITY.
FT REPEAT 2096 2140 BY SIMILARITY.
FT REPEAT 2141 2185 BY SIMILARITY.
FT REPEAT 2186 2230 BY SIMILARITY.
FT REPEAT 2231 2275 BY SIMILARITY.
FT REPEAT 2276 2320 BY SIMILARITY.
FT REPEAT 2321 2365 BY SIMILARITY.
FT REPEAT 2366 2410 BY SIMILARITY.
FT REPEAT 2411 2455 BY SIMILARITY.
FT REPEAT 2456 2500 BY SIMILARITY.
FT REPEAT 2501 2545 BY SIMILARITY.
FT REPEAT 2546 2590 BY SIMILARITY.
FT REPEAT 2591 2635 BY SIMILARITY.
FT REPEAT 2636 2680 BY SIMILARITY.
FT REPEAT 2681 2725 BY SIMILARITY.
FT REPEAT 2726 2770 BY SIMILARITY.
FT REPEAT 2771 2815 BY SIMILARITY.
FT REPEAT 2816 2860 BY SIMILARITY.
FT REPEAT 2861 2905 BY SIMILARITY.
FT REPEAT 2906 2950 BY SIMILARITY.
FT REPEAT 2951 2995 BY SIMILARITY.
FT REPEAT 2996 3040 BY SIMILARITY.
FT REPEAT 3041 3085 BY SIMILARITY.
FT REPEAT 3086 3130 BY SIMILARITY.
FT REPEAT 3131 3175 BY SIMILARITY.
FT REPEAT 3176 3220 BY SIMILARITY.
FT REPEAT 3221 3265 BY SIMILARITY.
FT REPEAT 3266 3310 BY SIMILARITY.
FT REPEAT 3311 3355 BY SIMILARITY.
FT REPEAT 3356 3400 BY SIMILARITY.
FT REPEAT 3401 3445 BY SIMILARITY.
FT REPEAT 3446 3490 BY SIMILARITY.
FT REPEAT 3491 3535 BY SIMILARITY.
FT REPEAT 3536 3580 BY SIMILARITY.
FT REPEAT 3581 3625 BY SIMILARITY.
FT REPEAT 3626 3670 BY SIMILARITY.
FT REPEAT 3671 3715 BY SIMILARITY.
FT REPEAT 3716 3760 BY SIMILARITY.
FT REPEAT 3761 3805 BY SIMILARITY.
FT REPEAT 3806 3850 BY SIMILARITY.
FT REPEAT 3851 3895 BY SIMILARITY.
FT REPEAT 3896 3940 BY SIMILARITY.
FT REPEAT 3941 3985 BY SIMILARITY.
FT REPEAT 3986 4030 BY SIMILARITY.
FT REPEAT 4031 4075 BY SIMILARITY.
FT REPEAT 4076 4120 BY SIMILARITY.
FT REPEAT 4121 4165 BY SIMILARITY.
FT REPEAT 4166 4210 BY SIMILARITY.
FT REPEAT 4211 4255 BY SIMILARITY.
FT REPEAT 4256 4300 BY SIMILARITY.
FT REPEAT 4301 4345 BY SIMILARITY.
FT REPEAT 4346 4390 BY SIMILARITY.
FT REPEAT 4391 4435 BY SIMILARITY.
FT REPEAT 4436 4480 BY SIMILARITY.
FT REPEAT 4481 4525 BY SIMILARITY.
FT REPEAT 4526 4570 BY SIMILARITY.
FT REPEAT 4571 4615 BY SIMILARITY.
FT REPEAT 4616 4660 BY SIMILARITY.
FT REPEAT 4661 4705 BY SIMILARITY.
FT REPEAT 4706 4750 BY SIMILARITY.
FT REPEAT 4751 4795 BY SIMILARITY.
FT REPEAT 4796 4840 BY SIMILARITY.
FT REPEAT 4841 4885 BY SIMILARITY.
FT REPEAT 4886 4930 BY SIMILARITY.
FT REPEAT 4931 4975 BY SIMILARITY.
FT REPEAT 4976 5020 BY SIMILARITY.
FT REPEAT 5021 5065 BY SIMILARITY.
FT REPEAT 5066 5110 BY SIMILARITY.
FT REPEAT 5111 5155 BY SIMILARITY.
FT REPEAT 5156 5200 BY SIMILARITY.
FT REPEAT 5201 5245 BY SIMILARITY.
FT REPEAT 5246 5290 BY SIMILARITY.
FT REPEAT 5291 5335 BY SIMILARITY.
FT REPEAT 5336 5380 BY SIMILARITY.
FT REPEAT 5381 5425 BY SIMILARITY.
FT REPEAT 5426 5470 BY SIMILARITY.
FT REPEAT 5471 5515 BY SIMILARITY.
FT REPEAT 5516 5560 BY SIMILARITY.
FT REPEAT 5561 5605 BY SIMILARITY.
FT REPEAT 5606 5650 BY SIMILARITY.
FT REPEAT 5651 5695 BY SIMILARITY.
FT REPEAT 5696 5740 BY SIMILARITY.
FT REPEAT 5741 5785 BY SIMILARITY.
FT REPEAT 5786 5830 BY SIMILARITY.
FT REPEAT 5831 5875 BY SIMILARITY.
FT REPEAT 5876 5920 BY SIMILARITY.
FT REPEAT 5921 5965 BY SIMILARITY.
FT REPEAT 5966 6010 BY SIMILARITY.
FT REPEAT 6011 6055 BY SIMILARITY.
FT REPEAT 6056 6100 BY SIMILARITY.
FT REPEAT 6101 6145 BY SIMILARITY.
FT REPEAT 6146 6190 BY SIMILARITY.
FT REPEAT 6191 6235 BY SIMILARITY.
FT REPEAT 6236 6280 BY SIMILARITY.
FT REPEAT 6281 6325 BY SIMILARITY.
FT REPEAT 6326 6370 BY SIMILARITY.
FT REPEAT 6371 6415 BY SIMILARITY.
FT REPEAT 6416 6460 BY SIMILARITY.
FT REPEAT 6461 6505 BY SIMILARITY.
FT REPEAT 6506 6550 BY SIMILARITY.
FT REPEAT 6551 6595 BY SIMILARITY.
FT REPEAT 6596 6640 BY SIMILARITY.
FT REPEAT 6641 6685 BY SIMILARITY.
FT REPEAT 6686 6730 BY SIMILARITY.
FT REPEAT 6731 6775 BY SIMILARITY.
FT REPEAT 6776 6820 BY SIMILARITY.
FT REPEAT 6821 6865 BY SIMILARITY.
FT REPEAT 6866 6910 BY SIMILARITY.
FT REPEAT 6911 6955 BY SIMILARITY.
FT REPEAT 6956 7000 BY SIMILARITY.
FT REPEAT 7001 7045 BY SIMILARITY.
FT REPEAT 7046 7090 BY SIMILARITY.
FT REPEAT 7091 7135 BY SIMILARITY.
FT REPEAT 7136 7180 BY SIMILARITY.
FT REPEAT 7181 7225 BY SIMILARITY.
FT REPEAT 7226 7270 BY SIMILARITY.
FT REPEAT 7271 7315 BY SIMILARITY.
FT REPEAT 7316 7360 BY SIMILARITY.
FT REPEAT 7361 7405 BY SIMILARITY.
FT REPEAT 7406 7450 BY SIMILARITY.
FT REPEAT 7451 7495 BY SIMILARITY.
FT REPEAT 7496 7540 BY SIMILARITY.
FT REPEAT 7541 7585 BY SIMILARITY.
FT REPEAT 7586 7630 BY SIMILARITY.
FT REPEAT 7631 7675 BY SIMILARITY.
FT REPEAT 7676 7720 BY SIMILARITY.
FT REPEAT 7721 7765 BY SIMILARITY.
FT REPEAT 7766 7810 BY SIMILARITY.
FT REPEAT 7811 7855 BY SIMILARITY.
FT REPEAT 7856 7900 BY SIMILARITY.
FT REPEAT 7901 7945 BY SIMILARITY.
FT REPEAT 7946 7990 BY SIMILARITY.
FT REPEAT 7991 8035 BY SIMILARITY.
FT REPEAT 8036 8080 BY SIMILARITY.
FT REPEAT 8081 8125 BY SIMILARITY.
FT REPEAT 8126 8170 BY SIMILARITY.
FT REPEAT 8171 8215 BY SIMILARITY.
FT REPEAT 8216 8260 BY SIMILARITY.
FT REPEAT 8261 8305 BY SIMILARITY.
FT REPEAT 8306 8350 BY SIMILARITY.
FT REPEAT 8351 8395 BY SIMILARITY.
FT REPEAT 8396 8440 BY SIMILARITY.
FT REPEAT 8441 8485 BY SIMILARITY.
FT REPEAT 8486 8530 BY SIMILARITY.
FT REPEAT 8531 8575 BY SIMILARITY.
FT REPEAT 8576 8620 BY SIMILARITY.
FT REPEAT 8621 8665 BY SIMILARITY.
FT REPEAT 8666 8710 BY SIMILARITY.
FT REPEAT 8711 8755 BY SIMILARITY.
FT REPEAT 8756 8800 BY SIMILARITY.
FT REPEAT 8801 8845 BY SIMILARITY.
FT REPEAT 8846 8890 BY SIMILARITY.
FT REPEAT 8891 8935 BY SIMILARITY.
FT REPEAT 8936 8980 BY SIMILARITY.
FT REPEAT 8981 9025 BY SIMILARITY.
FT REPEAT 9026 9070 BY SIMILARITY.
FT REPEAT 9071 9115 BY SIMILARITY.
FT REPEAT 9116 9160 BY SIMILARITY.
FT REPEAT 9161 9205 BY SIMILARITY.
FT REPEAT 9206 9250 BY SIMILARITY.
FT REPEAT 9251 9295 BY SIMILARITY.
FT REPEAT 9296 9340 BY SIMILARITY.
FT REPEAT 9341 9385 BY SIMILARITY.
FT REPEAT 9386 9430 BY SIMILARITY.
FT REPEAT 9431 9475 BY SIMILARITY.
FT REPEAT 9476 9520 BY SIMILARITY.
FT REPEAT 9521 9565 BY SIMILARITY.
FT REPEAT 9566 9610 BY SIMILARITY.
FT REPEAT 9611 9655 BY SIMILARITY.
FT REPEAT 9656 9700 BY SIMILARITY.
FT REPEAT 9701 9745 BY SIMILARITY.
FT REPEAT 9746 9790 BY SIMILARITY.
FT REPEAT 9791 9835 BY SIMILARITY.
FT REPEAT 9836 9880 BY SIMILARITY.
FT REPEAT 9881 9925 BY SIMILARITY.
FT REPEAT 9926 9970 BY SIMILARITY.
FT REPEAT 9971 10015 BY SIMILARITY.
FT REPEAT 10016 10060 BY SIMILARITY.
FT REPEAT 10061 10105 BY SIMILARITY.
FT REPEAT 10106 10150 BY SIMILARITY.
FT REPEAT 10151 10195 BY SIMILARITY.
FT REPEAT 10196 10240 BY SIMILARITY.
FT REPEAT 10241 10285 BY SIMILARITY.
FT REPEAT 10286 10330 BY SIMILARITY.
FT REPEAT 10331 10375 BY SIMILARITY.
FT REPEAT 10376 10420 BY SIMILARITY.
FT REPEAT 10421 10465 BY SIMILARITY.
FT REPEAT 10466 10510 BY SIMILARITY.
FT REPEAT 10511 10555 BY SIMILARITY.
FT REPEAT 10556 10600 BY SIMILARITY.
FT REPEAT 10601 10645 BY SIMILARITY.
FT REPEAT 10646 10690 BY SIMILARITY.
FT REPEAT 10691 10735 BY SIMILARITY.
FT REPEAT 10736 10780 BY SIMILARITY.
FT REPEAT 10781 10825 BY SIMILARITY.
FT REPEAT 10826 10870 BY SIMILARITY.
FT REPEAT 10871 10915 BY SIMILARITY.
FT REPEAT 10916 10960 BY SIMILARITY.
FT REPEAT 10961 11005 BY SIMILARITY.
FT REPEAT 11006 11050 BY SIMILARITY.
FT REPEAT 11051 11095 BY SIMILARITY.
FT REPEAT 11096 11140 BY SIMILARITY.
FT REPEAT 11141 11185 BY SIMILARITY.
FT REPEAT 11186 11230 BY SIMILARITY.
FT REPEAT 11231 11275 BY SIMILARITY.
FT REPEAT 11276 11320 BY SIMILARITY.
FT REPEAT 11321 11365 BY SIMILARITY.
FT REPEAT 11366 11410 BY SIMILARITY.
FT REPEAT 11411 11455 BY SIMILARITY.
FT REPEAT 11456 11500 BY SIMILARITY.
FT REPEAT 11501 11545 BY SIMILARITY.
FT REPEAT 11546 11590 BY SIMILARITY.
FT REPEAT 11591 11635 BY SIMILARITY.
FT REPEAT 11636 11680 BY SIMILARITY.
FT REPEAT 11681 11725 BY SIMILARITY.
FT REPEAT 11726 11770 BY SIMILARITY.
FT REPEAT 11771 11815 BY SIMILARITY.
FT REPEAT 11816 11860 BY SIMILARITY.
FT REPEAT 11861 11905 BY SIMILARITY.
FT REPEAT 11906 11950 BY SIMILARITY.
FT REPEAT 11951 11995 BY SIMILARITY.
FT REPEAT 11996 12040 BY SIMILARITY.
FT REPEAT 12041 12085 BY SIMILARITY.
FT REPEAT 12086 12130 BY SIMILARITY.
FT REPEAT 12131 12175 BY SIMILARITY.
FT REPEAT 12176 12220 BY SIMILARITY.
FT REPEAT 12221 12265 BY SIMILARITY.
FT REPEAT 12266 12310 BY SIMILARITY.
FT REPEAT 12311 12355 BY SIMILARITY.
FT REPEAT 12356 12400 BY SIMILARITY.
FT REPEAT 12401 12445 BY SIMILARITY.
FT REPEAT 12446 12490 BY SIMILARITY.
FT REPEAT 12491 12535 BY SIMILARITY.
FT REPEAT 12536 12580 BY SIMILARITY.
FT REPEAT 12581 12625 BY SIMILARITY.
FT REPEAT 12626 12670 BY SIMILARITY.
FT REPEAT 12671 12715 BY SIMILARITY.
FT REPEAT 12716 12760 BY SIMILARITY.
FT REPEAT 12761 12805 BY SIMILARITY.
FT REPEAT 12806 12850 BY SIMILARITY.
FT REPEAT 12851 12895 BY SIMILARITY.
FT REPEAT 12896 12940 BY SIMILARITY.
FT REPEAT 12941 12985 BY SIMILARITY.
FT REPEAT 12986 13030 BY SIMILARITY.
FT REPEAT 13031 13075 BY SIMILARITY.
FT REPEAT 13076 13120 BY SIMILARITY.
FT REPEAT 13121 13165 BY SIMILARITY.
FT REPEAT 13166 13210 BY SIMILARITY.
FT REPEAT 13211 13255 BY SIMILARITY.
FT REPEAT 13256 13300 BY SIMILARITY.
FT REPEAT 13301 13345 BY SIMILARITY.
FT REPEAT 13346 13390 BY SIMILARITY.
FT REPEAT 13391 13435 BY SIMILARITY.
FT REPEAT 13436 13480 BY SIMILARITY.
FT REPEAT 13481 13525 BY SIMILARITY.
FT REPEAT 13526 13570 BY SIMILARITY.
FT REPEAT 13571 13615 BY SIMILARITY.
FT REPEAT 13616 13660 BY SIMILARITY.
FT REPEAT 13661 13705 BY SIMILARITY.
FT REPEAT 13706 13750 BY SIMILARITY.
FT REPEAT 13751 13795 BY SIMILARITY.
FT REPEAT 13796 13840 BY SIMILARITY.
FT REPEAT 13841 13885 BY SIMILARITY.
FT REPEAT 13886 13930 BY SIMILARITY.
FT REPEAT 13931 13975 BY SIMILARITY.
FT REPEAT 13976 14020 BY SIMILARITY.
FT REPEAT 14021 14065 BY SIMILARITY.
FT REPEAT 14066 14110 BY SIMILARITY.
FT REPEAT 14111 14155 BY SIMILARITY.
FT REPEAT 14156 14200 BY SIMILARITY.
FT REPEAT 14201 14245 BY SIMILARITY.
FT REPEAT 14246 14290 BY SIMILARITY.
FT REPEAT 14291 14335 BY SIMILARITY.
FT REPEAT 14336 14380 BY SIMILARITY.
FT REPEAT 14381 14425 BY SIMILARITY.
FT REPEAT 14426 14470 BY SIMILARITY.
FT REPEAT 14471 14515 BY SIMILARITY.
FT REPEAT 14516 14560 BY SIMILARITY.
FT REPEAT 14561 14605 BY SIMILARITY.
FT REPEAT 14606 14650 BY SIMILARITY.
FT REPEAT 14651 14695 BY SIMILARITY.
FT REPEAT 14696 14740 BY SIMILARITY.
FT REPEAT 14741 14785 BY SIMILARITY.
FT REPEAT 14786 14830 BY SIMILARITY.
FT REPEAT 14831 14875 BY SIMILARITY.
FT REPEAT 14876 14920 BY SIMILARITY.
FT REPEAT 14921 14965 BY SIMILARITY.
FT REPEAT 14966 15010 BY SIMILARITY.
FT REPEAT 15011 15055 BY SIMILARITY.
FT REPEAT 15056 15100 BY SIMILARITY.
FT REPEAT 15101 15145 BY SIMILARITY.
FT REPEAT 15146 15190 BY SIMILARITY.
FT REPEAT 15191 15235 BY SIMILARITY.
FT REPEAT 15236 15280 BY SIMILARITY.
FT REPEAT 15281 15325 BY SIMILARITY.
FT REPEAT 15326 15370 BY SIMILARITY.
FT REPEAT 15371 15415 BY SIMILARITY.
FT REPEAT 15416 15460 BY SIMILARITY.
FT REPEAT 15461 15505 BY SIMILARITY.
FT REPEAT 15506 15550 BY SIMILARITY.
FT REPEAT 15551 15595 BY SIMILARITY.
FT REPEAT 15596 15640 BY SIMILARITY.
FT REPEAT 15641 15685 BY SIMILARITY.
FT REPEAT 15686 15730 BY SIMILARITY.
FT REPEAT 15731 15775 BY SIMILARITY.
FT REPEAT 15776 15820 BY SIMILARITY.
FT REPEAT 15821 15865 BY SIMILARITY.
FT REPEAT 15866 15910 BY SIMILARITY.
FT REPEAT 15911 15955 BY SIMILARITY.
FT REPEAT 15956 16000 BY SIMILARITY.
FT REPEAT 16001 16045 BY SIMILARITY.
FT REPEAT 16046 16090 BY SIMILARITY.
FT REPEAT 16091 16135 BY SIMILARITY.
FT REPEAT 16136 16180 BY SIMILARITY.
FT REPEAT 16181 16225 BY SIMILARITY.
FT REPEAT 16226 16270 BY SIMILARITY.
FT REPEAT 16271 16315 BY SIMILARITY.
FT REPEAT 16316 16360 BY SIMILARITY.
FT REPEAT 16361 16405 BY SIMILARITY.
FT REPEAT 16406 16450 BY SIMILARITY.
FT REPEAT 16451 16495 BY SIMILARITY.
FT REPEAT 16496 16540 BY SIMILARITY.
FT REPEAT 16541 16585 BY SIMILARITY.
FT REPEAT 16586 16630 BY SIMILARITY.
FT REPEAT 16631 16675 BY SIMILARITY.
FT REPEAT 16676 16720 BY SIMILARITY.
FT REPEAT 16721 16765 BY SIMILARITY.
FT REPEAT 16766 16810 BY SIMILARITY.
FT REPEAT 16811 16855 BY SIMILARITY.
FT REPEAT 16856 16900 BY SIMILARITY.
FT REPEAT 16901 16945 BY SIMILARITY.
FT REPEAT 16946 16990 BY SIMILARITY.
FT REPEAT 16991 17035 BY SIMILARITY.
FT REPEAT 17036 17080 BY SIMILARITY.
FT REPEAT 17081 17125 BY SIMILARITY.
FT REPEAT 17126 17170 BY SIMILARITY.
FT REPEAT 17171 17215 BY SIMILARITY.
FT REPEAT 17216 17260 BY SIMILARITY.
FT REPEAT 17261 17305 BY SIMILARITY.
FT REPEAT 17306 17350 BY SIMILARITY.
FT REPEAT 17351 17395 BY SIMILARITY.
FT REPEAT 17396 17440 BY SIMILARITY.
FT REPEAT 17441 17485 BY SIMILARITY.
FT REPEAT 17486 17530 BY SIMILARITY.
FT REPEAT 17531 17575 BY SIMILARITY.
FT REPEAT 17576 17620 BY SIMILARITY.
FT REPEAT 17621 17665 BY SIMILARITY.
FT REPEAT 17666 17710 BY SIMILARITY.
FT REPEAT 17711 17755 BY SIMILARITY.
FT REPEAT 17756 17800 BY SIMILARITY.
FT REPEAT 17801 17845 BY SIMILARITY.
FT REPEAT 17846 17890 BY SIMILARITY.
FT REPEAT 17891 17935 BY SIMILARITY.
FT REPEAT 17936 17980 BY SIMILARITY.
FT REPEAT 17981 18025 BY SIMILARITY.
FT REPEAT 18026 18070 BY SIMILARITY.
FT REPEAT 18071 18115 BY SIMILARITY.
FT REPEAT 18116 18160 BY SIMILARITY.
FT REPEAT 18161 18205 BY SIMILARITY.
FT REPEAT 18206 18250 BY SIMILARITY.
FT REPEAT 18251 18295 BY SIMILARITY.
FT REPEAT 18296 18340 BY SIMILARITY.
FT REPEAT 18341 18385 BY SIMILARITY.
FT REPEAT 18386 18430 BY SIMILARITY.
FT REPEAT 18431 18475 BY SIMILARITY.
FT REPEAT 18476 18520 BY SIMILARITY.
FT REPEAT 18521 18565 BY SIMILARITY.
FT REPEAT 18566 18610 BY SIMILARITY.
FT REPEAT 18611 18655 BY SIMILARITY.
FT REPEAT 18656 18700 BY SIMILARITY.
FT REPEAT 18701 18745 BY SIMILARITY.
FT REPEAT 18746 18790 BY SIMILARITY.
FT REPEAT 18791 18835 BY SIMILARITY.
FT REPEAT 18836 18880 BY SIMILARITY.
FT REPEAT 18881 18925 BY SIMILARITY.
FT REPEAT 18926 18970 BY SIMILARITY.
FT REPEAT 18971 19015 BY SIMILARITY.
FT REPEAT 19016 19060 BY SIMILARITY.
FT REPEAT 19061 19105 BY SIMILARITY.
FT REPEAT 19106 19150 BY SIMILARITY.
FT REPEAT 19151 19195 BY SIMILARITY.
FT REPEAT 19196 19240 BY SIMILARITY.
FT REPEAT 19241 19285 BY SIMILARITY.
FT REPEAT 19286 19330 BY SIMILARITY.
FT REPEAT 19331 19375 BY SIMILARITY.
FT REPEAT 19376 19420 BY SIMILARITY.
FT REPEAT 19421 19465 BY SIMILARITY.
FT REPEAT 19466 19510 BY SIMILARITY.
FT REPEAT 19511 19555 BY SIMILARITY.
FT REPEAT 19556 19600 BY SIMILARITY.
FT REPEAT 19601 19645 BY SIMILARITY.
FT REPEAT 19646 19690 BY SIMILARITY.
FT REPEAT 19691 19735 BY SIMILARITY.
FT REPEAT 19736 19780 BY SIMILARITY.
FT REPEAT 19781 19825 BY SIMILARITY.
FT REPEAT 19826 19870 BY SIMILARITY.
FT REPEAT 19871 19915 BY SIMILARITY.
FT REPEAT 19916 19960 BY SIMILARITY.
FT REPEAT 19961 20005 BY SIMILARITY.
FT REPEAT 20006 20050 BY SIMILARITY.
FT REPEAT 20051 20095 BY SIMILARITY.
FT REPEAT 20096 20140 BY SIMILARITY.
FT REPEAT 20141 20185 BY SIMILARITY.
FT REPEAT 20186 20230 BY SIMILARITY.
FT REPEAT 20231 20275 BY SIMILARITY.
FT REPEAT 20276 20320 BY SIMILARITY.
FT REPEAT 20321 20365 BY SIMILARITY.
FT REPEAT 20366 20410 BY SIMILARITY.
FT REPEAT 20411 20455 BY SIMILARITY.
FT REPEAT 20456 20500 BY SIMILARITY.
FT REPEAT 20501 20545 BY SIMILARITY.
FT REPEAT 20546 20590 BY SIMILARITY.
FT REPEAT 20591 20635 BY SIMILARITY.
FT REPEAT 20636 20680 BY SIMILARITY.
FT REPEAT 20681 20725 BY SIMILARITY.
FT REPEAT 20726 20770 BY SIMILARITY.
FT REPEAT 20771 20815 BY SIMILARITY.
FT REPEAT 20816 20860 BY SIMILARITY.
FT REPEAT 20861 20905 BY SIMILARITY.
FT REPEAT 20906 20950 BY SIMILARITY.
FT REPEAT 20951 20995 BY SIMILARITY.
FT REPEAT 20996 21040 BY SIMILARITY.
FT REPEAT 21041 21085 BY SIMILARITY.
FT REPEAT 21086 21130 BY SIMILARITY.
FT REPEAT 21131 21175 BY SIMILARITY.
FT REPEAT 21176 21220 BY SIMILARITY.
FT REPEAT 21221 21265 BY SIMILARITY.
FT REPEAT 21266 21310 BY SIMILARITY.
FT REPEAT 21311 21355 BY SIMILARITY.
FT REPEAT 21356 21400 BY SIMILARITY.
FT REPEAT 21401 21445 BY SIMILARITY.
FT REPEAT 21446 21490 BY SIMILARITY.
FT REPEAT 21491 21535 BY SIMILARITY.
FT REPEAT 21536 21580 BY SIMILARITY.
FT REPEAT 21581 21625 BY SIMILARITY.
FT REPEAT 21626 21670 BY SIMILARITY.
FT REPEAT 21671 21715 BY SIMILARITY.
FT REPEAT 21716 21760 BY SIMILARITY.
FT REPEAT 21761 21805 BY SIMILARITY.
FT REPEAT 21806 21850 BY SIMILARITY.
FT REPEAT 21851 21895 BY SIMILARITY.
FT REPEAT 21896 21940 BY SIMILARITY.
FT REPEAT 21941 21985 BY SIMILARITY.
FT REPEAT 21986 22030 BY SIMILARITY.
FT REPEAT 22031 22075 BY SIMILARITY.
FT REPEAT 22076 22120 BY SIMILARITY.
FT REPEAT 22121 22165 BY SIMILARITY.
FT REPEAT 22166 22210 BY SIMILARITY.
FT REPEAT 22211 22255 BY SIMILARITY.
FT REPEAT 22256 22300 BY SIMILARITY.
FT REPEAT 22301 22345 BY SIMILARITY.
FT REPEAT 22346 22390 BY SIMILARITY.
FT REPEAT 22391 22435 BY SIMILARITY.
FT REPEAT 22436 22480 BY SIMILARITY.
FT REPEAT 22481 22525 BY SIMILARITY.
FT REPEAT 22526 22570 BY SIMILARITY.
FT REPEAT 22571 22615 BY SIMILARITY.
FT REPEAT 22616 22660 BY SIMILARITY.
FT REPEAT 22661 22705 BY SIMILARITY.
FT REPEAT 22706 22750 BY SIMILARITY.
FT REPEAT 22751 22795 BY SIMILARITY.
FT REPEAT 22796 22840 BY SIMILARITY.
FT REPEAT 22841 22885 BY SIMILARITY.
FT REPEAT 22886 22930 BY SIMILARITY.
FT REPEAT 22931 22975 BY SIMILARITY.
FT REPEAT 22976 23020 BY SIMILARITY.
FT REPEAT 23021 23065 BY SIMILARITY.
FT REPEAT 23066 23110 BY SIMILARITY.
FT REPEAT 23111 23155 BY SIMILARITY.
FT REPEAT 23
```

```
FT DISULFID 497 512 BY SIMILARITY.
FT DISULFID 514 523 BY SIMILARITY.
FT DISULFID 530 541 BY SIMILARITY.
FT DISULFID 535 550 BY SIMILARITY.
FT DISULFID 552 561 BY SIMILARITY.
FT DISULFID 568 579 BY SIMILARITY.
FT DISULFID 573 588 BY SIMILARITY.
FT DISULFID 590 599 BY SIMILARITY.
FT DISULFID 606 616 BY SIMILARITY.
FT DISULFID 611 625 BY SIMILARITY.
FT DISULFID 627 636 BY SIMILARITY.
FT DISULFID 643 654 BY SIMILARITY.
FT DISULFID 648 663 BY SIMILARITY.
FT DISULFID 665 674 BY SIMILARITY.
FT DISULFID 681 692 BY SIMILARITY.
```

Query Match 3 18; Score 124.5; DB 1; Length 2703;
Best Local Similarity 23.98; Pred. No. 1.2; Mismatches 115; Indels 77; Gaps 15;
Matches 75; Conservative 47;

```
QY 39 MESQFGEDRKFPQIRVNLNTRKG--TGASQDPNRFDRD-----LFNAVSRG-- 86
DB 1856 MVSEYEADQQRVMSQAHLDVVYRAIMTPRAHQDGSKHDVDARGPGCLTPMLAAVRGGG 1915
QY 87 --VPEDLAGLPEYLSKTSKYLD-----SEYTESGTGKTCIMKAVLNKQGVNACILP 137
DB 1916 LPTGEDIEENNED---STAQVISDLAOGAELNATMDKTGETSLHLARFARADAKRL- 1971
QY 138 LQIDRDGNGPQPLVNAOCTDDYRGHSALHITAIEKRSLOCVKLIYENGA-NVHARACGR 196
DB 1972 -----DAG-----ADANCODN--TGRTPHLAAVNAADAGVFQILLNRNTNLNR---- 2014
QY 197 FPOKGGCTCFYEGELPLSLAACKOMDVSYLLENPHOPASLOATDSOGNTVLLHALMIS 256
DB 2015 -----MHDGTPLLIARLAIEGWEDLIT---ADADRMAADNSGKTALHMAAAYN 2062
QY 257 DNSAENIALVTSMYDGLQAGARLCPTVQLEDIRNLQDLTPKLAKESGKIETFRHIL-- 314
DB 2063 NTEAVNI-----LMHMHAN-----RDADCKDETPFLAAREGSYEACKALLDN 2106
QY 315 --QREFSGLSHLR 326
DB 2107 FANREIT--DHMDR 2118
```

Search completed: July 18, 2001, 16:00:28
Job time: 131 sec